

Serial Number: 09/645,706

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

---

### In the Claims

1. (Currently amended) A synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a polypeptide, wherein the codon composition and transcription regulatory sequence composition of the synthetic nucleic acid molecule is different than that of a wild type nucleic acid sequence encoding a polypeptide which has at least 85% sequence identity to the polypeptide encoded by the synthetic nucleic acid molecule, wherein the codons which differ are selected so as to result in the synthetic nucleic acid molecule having a codon composition differing at more than 25% of the codons from a the wild type nucleic acid sequence encoding a polypeptide, and having a reduced number of transcription factor binding sequences at least 3-fold fewer transcription regulatory sequences and optionally a reduced number of relative to the average number of such sequences resulting from random selections of codons at the codons which differ, wherein the transcription regulatory sequences are selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A) addition sites and promoter sequences, and wherein the polypeptide encoded by the synthetic nucleic acid molecule has at least 85% sequence identity to the polypeptide encoded by the wild type nucleic acid sequence; and wherein the synthetic nucleic acid molecule has reduced aberrant transcription relative to the transcription of the wild type nucleic acid sequence.
2. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has at least 5-fold fewer transcription regulatory sequences.
3. (Original) The synthetic nucleic acid molecule of claim 1 wherein the codon composition of the synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 35% of the codons.

4. (Original) The synthetic nucleic acid molecule of claim 1 wherein the codon composition of the synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 45% of the codons.
5. (Original) The synthetic nucleic acid molecule of claim 1 wherein the codon composition of the synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 55% of the codons.
6. (Original) The synthetic nucleic acid molecule of claim 1 wherein the majority of codons which differ are ones that are preferred codons of a desired host cell.
7. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule encodes a reporter molecule.
8. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule encodes a selectable marker protein.
9. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule encodes a luciferase.
10. (Canceled)
11. (Original) The synthetic nucleic acid molecule of claim 9 wherein the wild type nucleic acid sequence encodes a beetle luciferase.
12. (Original) The synthetic nucleic acid molecule of claim 11 wherein the synthetic nucleic acid molecule encodes the amino acid valine at position 224.
13. (Canceled)

14. (Original) The synthetic nucleic acid molecule of claim 1 wherein the majority of codons which differ in the synthetic nucleic acid molecule are those which are employed more frequently in mammals.

15. (Original) The synthetic nucleic acid molecule of claim 1 wherein the majority of codons which differ in the synthetic nucleic acid molecule are those which are preferred codons in humans.

16-17. (Canceled)

18. (Original) The synthetic nucleic acid molecule of claim 9 wherein the synthetic nucleic acid molecule comprises SEQ ID NO:7 (GRver5), SEQ ID NO:8 (GRver6), SEQ ID NO:9 (GRver5.1), or SEQ ID NO:297 (GRver5.1).

19. (Canceled)

20. (Original) The synthetic nucleic acid molecule of claim 15 wherein the majority of codons which differ are the human codons CGC, CTG, TCT, AGC, ACC, CCA, CCT, GCC, GGC, GTG, ATC, ATT, AAG, AAC, CAG, CAC, GAG, GAC, TAC, TGC and TTC.

21. (Original) The synthetic nucleic acid molecule of claim 15 wherein the majority of codons which differ are the human codons CGC, CTG, TCT, ACC, CCA, GCC, GGC, GTC, and ATC or codons CGT, TTG, AGC, ACT, CCT, GCT, GGT, GTG and ATT.

22-23. (Canceled)

24. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule is expressed in a mammalian host cell at a level which is greater than that of the wild type nucleic acid sequence.
25. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of CTG or TTG leucine-encoding codons.
26. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of GTG or GTC valine-encoding codons.
27. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of GGC or GGT glycine-encoding codons.
28. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule an increased number of ATC or ATT isoleucine-encoding codons.
29. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of CCA or CCT proline-encoding codons.
30. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of CGC or CGT arginine-encoding codons.

Serial Number: 09/645,706

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

---

31. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of AGC or TCT serine-encoding codons.
32. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of ACC or ACT threonine-encoding codons.
33. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule has an increased number of GCC or GCT alanine-encoding codons.
34. (Original) The synthetic nucleic acid molecule of claim 1 wherein the codons in the synthetic nucleic acid molecule which differ encode the same amino acids as the corresponding codons in the wild type nucleic acid sequence.
35. (Original) A plasmid comprising the synthetic nucleic acid molecule of claim 1.
36. (Original) An expression vector comprising the synthetic nucleic acid molecule of claim 1 linked to a promoter functional in a cell.
37. (Original) The expression vector of claim 36 wherein the synthetic nucleic acid molecule is operatively linked to a Kozak consensus sequence.
38. (Original) The expression vector of claim 36 wherein the promoter is functional in a mammalian cell.
39. (Original) The expression vector of claim 36 wherein the promoter is functional in a human cell.

40. (Canceled)
41. (Original) The expression vector of claim 36 wherein the expression vector further comprises a multiple cloning site.
42. (Original) The expression vector of claim 41 wherein the expression vector comprises a multiple cloning site positioned between the promoter and the synthetic nucleic acid molecule.
43. (Original) The expression vector of claim 41 wherein the expression vector comprises a multiple cloning site positioned downstream from the synthetic nucleic acid molecule.
44. (Original) A host cell comprising the expression vector of claim 36.
45. (Currently amended) A ~~reporter gene expression~~ kit comprising, in suitable container means, the expression vector of claim 36, wherein the synthetic nucleic acid molecule encodes a reporter molecule.
46. (Canceled)
47. (Currently amended) A polynucleotide which hybridizes under stringent medium stringency hybridization conditions to SEQ ID NO:22 (Rluc-final), SEQ ID NO:9 (GRver5.1), SEQ ID NO:18 (RD156-1H9), SEQ ID NO:297 (GRver5.1), SEQ ID NO:301 (RD156-1H9), or the complement thereof, wherein the codon composition and transcription regulatory sequence composition of the polynucleotide is different than that of a wild type nucleic acid sequence encoding a polypeptide which has at least 85% sequence identity to the polypeptide encoded by SEQ ID NO:22, SEQ ID NO:9, SEQ ID NO:18, SEQ ID NO:297, or SEQ ID NO:301, wherein the codons which differ are selected so as

to result in the polynucleotide having a codon composition differing at more than 25% of the codons from the wild type nucleic acid sequence and having a reduced number of transcription factor binding sequences and optionally a reduced number of sequences selected from the group consisting of intron splice sites, poly(A) addition sites and promoter sequences.

48-59. (Canceled)

60. (Original) The synthetic nucleic acid molecule of claim 1 wherein the synthetic nucleic acid molecule is expressed at a level which is at least 110% of that of the wild type nucleic acid sequence in a cell or cell extract under identical conditions.
61. (Original) The synthetic nucleic acid molecule of claim 1 wherein the polypeptide encoded by the synthetic nucleic acid molecule has at least 90% contiguous sequence identity to the polypeptide encoded by the wild type nucleic acid sequence.
62. (Original) The synthetic nucleic acid molecule of claim 1 wherein the polypeptide encoded by the synthetic nucleic acid molecule is identical in amino acid sequence to the polypeptide encoded by the wild type nucleic acid sequence.
63. (Currently amended) A vector comprising a gene of interest and backbone sequences, wherein the backbone sequences comprise a synthetic nucleic acid molecule having at least 3-fold fewer sequences selected from the group consisting of transcriptional regulatory sequences transcription factor binding sequences, relative to a vector comprising a parent nucleic acid sequence, wherein the transcription regulatory sequences are selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A)

addition sites and promoter sequences, relative to a vector comprising a parent nucleic acid sequence.

65. (Original) The vector of claim 63 wherein the synthetic nucleic acid molecule does not encode a polypeptide.

65-66. (Canceled)

67. (New) A synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a luciferase, wherein the codon composition and transcription regulatory sequence composition of the synthetic nucleic acid molecule is different than that of a wild type nucleic acid sequence encoding a luciferase which has at least 85% sequence identity to the luciferase encoded by the synthetic nucleic acid molecule, wherein the codons which differ are selected so as to result in a synthetic nucleic acid molecule having a codon composition differing at more than 25% of the codons from the wild type nucleic acid sequence and having a reduced number of transcription factor binding sequences and optionally sequences selected from the group consisting of intron splice sites, poly(A) addition sites and promoter sequences, wherein the synthetic nucleic acid molecule has reduced aberrant transcription relative to the transcription of the wild type nucleic acid sequence, and wherein the codons which differ are codons which are employed more frequently in mammals.
68. (New) The synthetic nucleic acid molecule of claim 1 or 67 which has at least 3-fold fewer sequences selected from the group consisting of transcription factor binding sequences, intron splice sites, poly(A) addition sites, and promoter sequences.